

~~RESTRICTED~~
~~SECURITY INFORMATION~~

40-1/6

NOTE ON INFORMATION ENTROPY
FOR
QUANTIZED NORMAL DISTRIBUTION

Report Number R-40

Copy ____ of 150 Copies

May 1953

Prepared by:

Albert A. Blank

CONTROL SYSTEMS LABORATORY
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS
Contract DA-11-022-ORD-721

Numbered Pages: 6

~~RESTRICTED~~
~~SECURITY INFORMATION~~
UNCLASSIFIED

THIS DOCUMENT CONTAINS INFORMATION AFFECTING THE NATIONAL
DEFENSE OF THE UNITED STATES WITHIN THE MEANING OF THE
ESPIONAGE LAWS, TITLE 18, U.S.C., SECTIONS 793 AND 794. ITS
TRANSMISSION OR THE REVELATION OF ITS CONTENTS IN ANY MANNER
TO AN UNAUTHORIZED PERSON IS PROHIBITED BY LAW.

UNCLASSIFIED

~~RESTRICTED~~
~~SECRET~~

Acq. # 2142 R/V

CONTROL SYSTEMS LABORATORY

NOTE ON INFORMATION ENTROPY
FOR
QUANTIZED NORMAL DISTRIBUTION

Report Number R-40

May 1953

Copy 13 of 150 Copies

Albert A. Blank

UNIVERSITY OF ILLINOIS · URBANA · ILLINOIS

UNCLASSIFIED

~~RESTRICTED~~
~~SECRET~~

RESTRICTED

40-2/6

FOREWORD

The purpose of this note, a purely utilitarian one, is to give a simple means of computing the Shannon-Wiener entropy function for a discrete statistics in those cases where it may safely be assumed that the statistics arise from the grouping of a normal probability distribution.

RESTRICTED

In information theory, the entropy function is computed differently for discrete and continuous distributions. In the discrete case, we define the information entropy as

$$(1) \quad H = - \sum p_n \log p_n$$

where p_n denotes the probability of an event in the n -th category. In the continuous case, we utilize the integral

$$(2) \quad H' = - \int_{\Omega} \log f(\xi) dF(\xi)$$

where $f(\xi)$ is the probability density of the random variable ξ , $F(\xi)$ the sum function and Ω is the sample space in which ξ varies. Now, H is never negative while H' may take on any value between plus and minus infinity. For example, in the case of the normal distribution, with probability density function

$$(3) \quad f(x) = e^{-\frac{x^2}{2\sigma^2}} / \sigma \sqrt{2\pi}$$

we have

$$(4) \quad H' = 1/2 \log 2\pi e\sigma^2$$

and since σ^2 may assume any positive value, H' may assume any real value whatever.*

*Shannon, C.E., and Weaver, W., "The Mathematical Theory of Communication", pp.54 ff., University of Illinois Press, Urbana, 1949.

Experimental data are generally obtained as discontinuous distributions; therefore, the appropriate entropy function to use is that of (1). It is often desirable to treat the situation as if the distribution were continuous, and use the function (2). In that case, there is a question as to how well the function H' approximates the function H . If a discrete distribution can be interpreted as a quantization of a normal distribution, then it is especially convenient to use (4). This is a valid approximation if the class interval is not too large with respect to σ . If the precision is high and most values are localized in one compartment, then H will be very close to zero but H' will be large in value and negative, and, therefore, the approximation cannot be used in this case.

It then becomes necessary to specify the range in which H' is a good approximation to H . If we let Δ be the size of the compartment and set $\lambda = \sigma/\Delta$ then we may compute H as a function of λ from (1) with the stipulation

$$p_n(\lambda) = \int_{(n-1/2)/\lambda}^{(n+1/2)/\lambda} \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} dt \quad (n = 0, \pm 1, \pm 2, \dots)$$

RESTRICTED

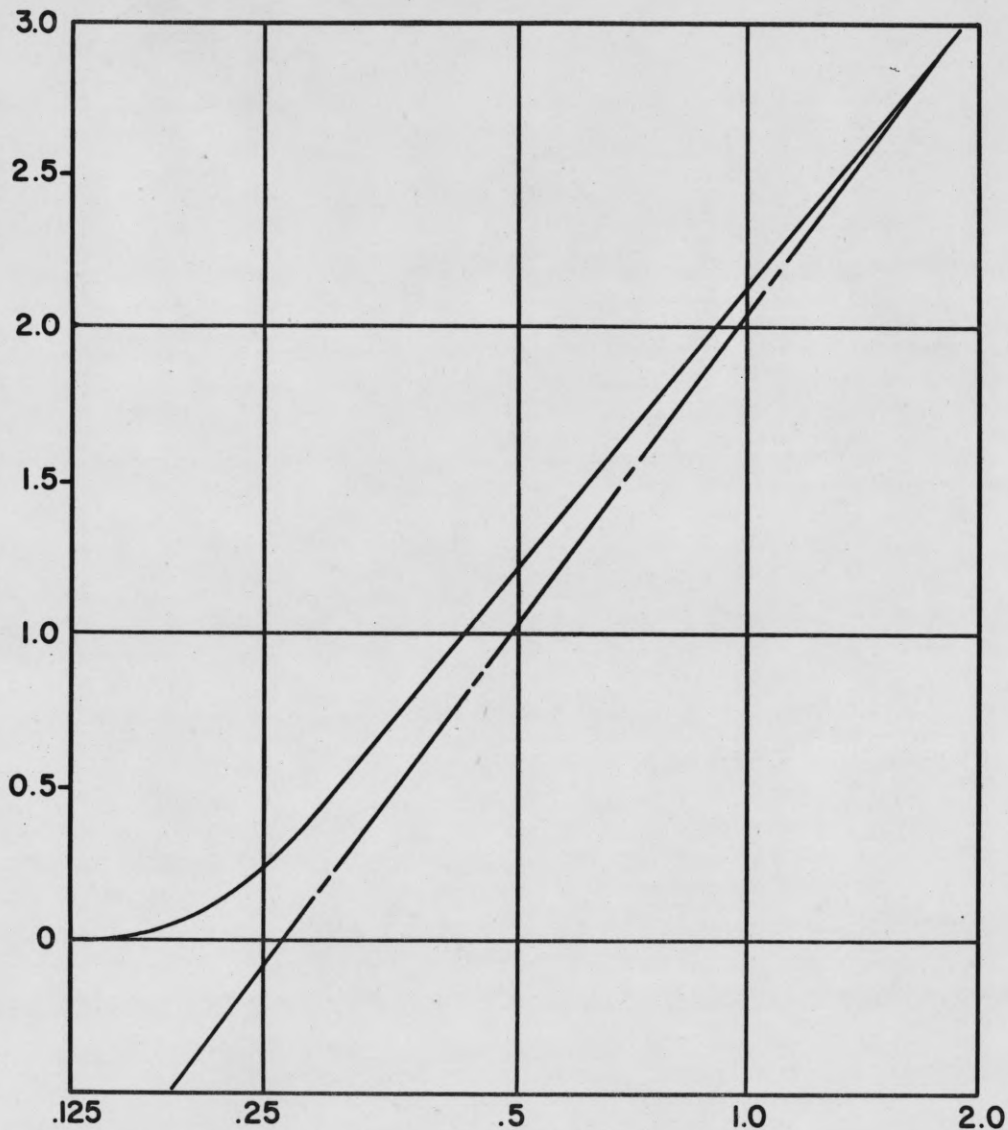
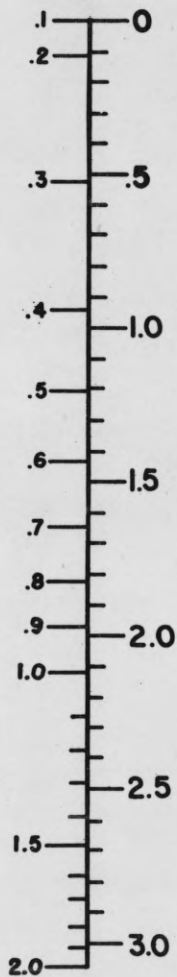
40-5/6

The curves $H'(\lambda)$ and $H(\lambda)$ are plotted together on the accompanying graph. For $\lambda > 2$, class interval $\Delta < \frac{\sigma}{2}$, we may for all practical purposes assume $H' = H$; for $\lambda < 0.1$, class interval $\Delta > 10\sigma$, we may assume $H = 0$. A nomogram for intermediate values of λ is included in the accompanying graph.

RESTRICTED

NOMOGRAM

$$\lambda = \frac{\sigma}{\Delta} H$$



$$H(\lambda) = - \sum_{-\infty}^{+\infty} p_n(\lambda) \log p_n(\lambda)$$

$$H'(\lambda) = \frac{1}{2} \log 2 \pi e \lambda^2$$

RESTRICTED

40-6/6